Discussion Papers No. 563, October 2008
Statistics Norway, Research Department

Chiara Pronzato

Why Educated Mothers don't make Educated Children?
A Statistical Study in the Intergenerational Transmission of Schooling

Abstract:
More educated parents are observed to have better educated children. From a policy point of view, however, it is important to distinguish between causation and selection. Researchers trying to control for unobserved ability have found conflicting results: in most cases, they have found a strong positive paternal effect but a negligible maternal effect. In this paper, I evaluate the impact on the robustness of the estimates of the characteristics of the samples commonly used in this strand of research: samples of small size, with low variability in parental education, not randomly selected from the population. The part of the educational distribution involved in any identification strategy seems to be a key aspect to take into account to reconcile previous results from the literature.

Keywords: intergenerational transmission, education, twin-estimator, sibling-estimator, power of the test

JEL classification: C23, I2

Acknowledgement: The Norwegian Research Council has provided financial support for this project. I am grateful to John Ermisch, Rolf Aaberge, Ugo Colombino, Helena Holmlund, Cheti Nicoletti, Richard Berthoud and Francesco Figari for their comments. Any error should be attributed to the author.

Address: Chiara Pronzato, DONDENA Centre for Research on Social Dynamics (Bocconi University), Milan (Italy). E-mail: chiara.pronzato@unibocconi.it
Other affilations: ISER (University of Essex), Statistics Norway, and CHILD.

Discussion Papers comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no
http://ideas.repec.org/s/ssb/dispap.html

For printed Discussion Papers contact:

Statistics Norway
Sales- and subscription service
NO-2225 Kongsvinger

Telephone: +47 62 88 55 00
Telefax: +47 62 88 55 95
E-mail: Salg-abonnement@ssb.no

1. Introduction

Much research, in recent years, has focused on the link between parental education and children's education. It has been shown empirically that more educated parents have better educated children. Using simple regression analysis, the correlation between parent's and child's education is strong and robust to a number of controls, sample selections, and countries (Haveman and Wolfe, 1995).

The policy implications of a causal link between parental education and children's education are huge. Increasing education today would lead to an increase in the schooling of the next generation and, in this way, to an improvement of later life outcomes such as health, wealth and productivity. From a policy point of view, however, it is important to distinguish between causation and selection. Better educated parents have, on average, higher abilities, which partially transmit to their children. Better schooling performances of their children could, theoretically, reflect this higher average ability.

When researchers have tried to control for unobserved ability, they have found conflicting results. In most cases, they have found a strong positive father's effect with a negligible mother's effect. In a few cases, a positive effect has been found for the mother and not for the father. These results depend on different identification strategies, and on different sources of information. In Section 2, I summarize techniques and results from previous studies, with particular attention to the work by Holmlund, Lindahl, and Plug (2008). Their study uses different identification strategies with the same data source concluding that the identification strategy matters.

The aim of this study is to explain the difference between mother's and father's effects on children's education from a statistical point of view. Why is it that women's education does not appear to raise the education of their children? In this paper, I evaluate the impact of selecting samples of small size, with low variability in parental education, and not always representative of the population, on the estimates of the intergenerational transmission of schooling.

In Section 3, I apply one identification strategy, the twin estimator, using Norwegian register data, and reach the same conclusion as others: a strong positive effect of father's education and no effect of mother's education. Then I show through simulations (Section 4) the extent to which small sample sizes and low variability in parental education is responsible for different results by gender, when using twins. In Section 5 I examine the effect of using selected samples of parents (at the bottom and at the top of the educational distribution) when analysing the intergenerational transmission of schooling. In Section 6 I redefine the outcome variable "years of schooling", exploiting more information available in the data, and replicate the analyses. Conclusions follow in Section 7.

2. The Intergenerational Transmission of Education in the Economic Literature

There are typically three strategies used in the literature to identify the effect of parental education on children's education: identical twins, adopted children, and using a reform of the educational system as an instrument for parental education.

Identical twins are the most similar individuals we can observe: they share the same family background, they experience lifetime events at the same time, and they share the same genes. When studying the intergenerational transmission of education, we compare the schooling of twins' children (i.e., cousins). These cousins share the ability and other family features transmitted[1] by the twins. On the other hand, they are exposed to different treatments: they can be the children of the more educated twin, the other parent (twin's spouse) has different characteristics, and some of them can grow up in a single parent household. In such studies, we can identify the effect of parental education on children's education, looking whether the child of the more educated twin has higher educational attainment than the child of the less educated twin. The shortcoming of this strategy is the assumption of random education between twins: why do children with identical abilities end up with different levels of education? If there are some characteristics which make one twin gain more education than the other, and if these characteristics can be transmitted to their children, then the resulting estimates are still biased. Despite this strong assumption, this method has been recognized as a good way to reduce ability bias (Bound and Solon, 1999). Another critique to the use of the twin-estimator derives by its sensitivity to measurement errors (Griliches, 1979; Ashenfelter and Krueger, 1994; Neumark, 1994; Bound and Solon, 1999; Light and Flores-Lagunes, 2006). From the point of view of external validity, we may wonder whether twins are representative of the whole population: we know, for example, that they have lower weight at birth, they are more likely to have problems in language, and they are brought up differently than other children (Mowrer, 1954; Mittler, 1971; Stewart, 2000; Schieve et al., 2002).

Another strategy used to eliminate, or at least reduce, the ability bias in the estimation of the intergenerational transmission of education is to use a sample of adopted children. In fact, there is no transmission of ability between adopted children and their adoptive parents. In this case, a relationship between parental and children's education should reveal a causal link between the two, while the comparison with estimates obtained from own-born children can be suggestive of the importance of the ability bias. The most common criticism of this strategy is that children are not

---

[1] Throughout the paper the expression ability comprises concepts like cleverness but also other family characteristics (non-cognitive skills, taste for education) which are shared by the twins, may influence schooling attainment and may be transmitted from parents to children.

randomly assigned: if adoption authorities have information on children's natural parents, they can use it to match children to adoptive couples. Another criticism derives from the fact that adoptive parents are not representative: they are, on average, older, better educated and more motivated than the overall population. This could threaten the generalization of the results.

Finally, other studies have used a reform of the educational system as an instrument for parental education. For example, increasing the age of compulsory schooling lengthens the years of education exogenously. Exploiting this exogenous variation, we can observe whether the children of parents whose education has increased because of the reform achieve more in the education system than children of parents not influenced by the reform. While there is no risk of ability bias, the ability of this approach to identify the causal effect depends on the strength of the instrument in generating enough variability. From an external validity point of view, results are difficult to generalize since the reform of compulsory schooling usually involves only the bottom part of the education distribution.

There are a number of papers which use different strategies, data, outcomes and control variables. I focus on those with "years of schooling" as the outcome variable and which take into account the level of education of both parents simultaneously. These studies are most comparable with the work carried out in this paper. Since the aim is to explain variation in the statistical significance of results, I indicate that a parameter estimate is significant at 10% level with 1 star, at 5% with 2 stars, at 1% with 3 stars.

The first work making use of twins in the study of the intergenerational transmission of education was published by Behrman and Rosenzweig (2002). They use a sample of monozygotic twins drawn from the Minnesota Twin Register. Information was obtained through a mail survey. They have 424 twin-mothers and 244 twin-fathers. They find a positive effect of father's schooling on children's years of schooling (0.340**) and a weak negative effect of mother's education (-0.263*). They suggest that this pattern is consistent with the fact that women's time is an important determinant of children's outcomes: the potential positive effect of mother's education is offset by the fact that more educated women spend more time in the labour market and less time with their children. Antonovics and Goldberger (2005) cast doubt on the construction of the dataset used by Behrman and Rosenzweig. After what they consider to be appropriate cleaning of the dataset, they have a sample of 90 twin-mothers and 47 twin-fathers. But the striking difference between genders remains: the effect of father's education is positive (0.477 – standard error not available) and mother's effect very close to zero (-0.003 – standard error not available).

To estimate the effect of mother's schooling on children's schooling, Plug (2004) controls for the effect of unobserved inherited abilities, obtaining identification by using adopted children instead of twins. He uses information collected in 1992 from 610 students who graduated from high schools in Wisconsin in 1957 and also finds a strong and positive effect of father's education (0.209***) and little effect of mother's education (0.089). He proposes different explanations for this result: some are substantive (better educated women spend less time with their children, differences in upbringing between own children and adopted children, adopted children different from other children in ways related to maternal schooling effects); some are more related to the design of the analysis (measurement error, heterogeneity with respect to the age adopted children enter their adoptive families, selection of highly educated mothers and consequently little variance in their education). Using Swedish adoption data, Björklund, Lindahl, and Plug (2006) are not only able to remove the "family fixed effect" from the effect of interest but also able to distinguish between pre-birth factors (genetics and prenatal environment) and post-birth environmental factors. They exploit the fact that Swedish register data contains information for both biological and adoptive parents. They have information on both pairs of parents for 2,125 adopted children. The effect of adoptive father's education is positive (0.094***) while that of adoptive mother's is small and insignificant (0.021).

Black, Devereux, and Salvanes (2005a) use a reform of compulsory schooling in Norway in the years 1947-1958 as an instrument for parental education. This reform resulted in 2 years more of schooling (from 7 years to 9 years). They use administrative data linked with the municipalities which implemented the reform, year by year. They find a significant effect of parental schooling only when selecting low educated parents (the ones most likely influenced by the reform): a positive effect of mother's education (0.122***) but a marginal effect of father's education (0.041).

These conflicting results obtained using various identification strategies and different datasets raise the question of what drives the differences. Are they data specific or do they depend on identification strategy? Holmlund, Lindahl, and Plug (2008) use different identification strategies with the same source of data, reaching the conclusion that it is identification that matters. They select from Swedish register data parents born between 1945 and 1955, whose experience of the reform of compulsory schooling depended on the municipality of residence. They find a positive result of mother's education (0.150*) when selecting those with low education, while they do not find any significant corresponding result for father's education (-0.030). Using information from 192 Swedish children adopted by parents of the 1945-1955 birth-cohorts, they find a positive effect of father's education (0.130*) and no effect of mother's education (-0.022). Finally, they have 3,850 twin-mothers and 2,306 twin-fathers (only half of them monozygotic), from which they estimate a mother's effect equal to 0.034 and a father's effect equal to 0.164***.

The aim of this paper is to show to what extent samples of small size, with low variability in parental education, not randomly selected from the population, can affect the estimates of the intergenerational schooling transmission.

3. Replicated Results

The first step of my empirical research is to replicate the analyses in previous work (Behrman and Rosenzweig, 2002; Antonovics and Goldberger, 2005; Holmlund, Lindahl, and Plug, 2008) in order to outline similarities and differences in the results and in the definition of the variables. The intergenerational schooling effect is estimated by using twins. The twin-estimator indicates whether the child of the more educated twin obtains more schooling qualifications, controlling for the ability transmitted by the parent.

I define *Y* the child's years of schooling, *X* the parent's years of schooling, and *Z* other factors which may influence the child's education. For the child *i* in the family *j*, we have

$$(1) \qquad y_{ji} = \beta x_{ji} + z'_{ji}\alpha + u_j + \varepsilon_{ji}$$

where β is the effect of parental education on the child's years of schooling, α the effect of other factors, u is the level of ability shared in the family *j*, and ε is the error term assumed to be white noise. A pooled regression of *Y* on *X* and *Z* (cross section) is not appropriate since it ignores the ability *u* shared in each family, which may be correlated with parental education. We can eliminate *u* from the equation, differencing the data in the following way

$$(2) \qquad \left(y_{ji} - \overline{y}_j\right) = \beta\left(x_{ji} - \overline{x}_j\right) + \left(z_{ji} - \overline{z}_j\right)'\alpha + \left(\varepsilon_{ji} - \overline{\varepsilon}_j\right)$$

where $\overline{y}_j$ is the average years of schooling of children in family *j*, $\overline{x}_j$ is the average years of schooling of twins in family *j*, and $\overline{z}_j$ are other average characteristics of the family *j*. The above regression is the basis of the twin-estimator in this empirical application.

The informational basis for the empirical analysis is a register household panel data set covering the entire resident population of Norway for the years 1993-2001. In this dataset, I have information on household size and composition as well as individual information such as place of residence, date of birth, educational level, etc. Concerning measurement error, register data have the advantage of reporting information from what people do, and not from what people say.

Twins are defined as people of the same sex, born from the same parents, in the same calendar year and month. Around 50% are likely to be identical twins (monozygotic), while the other 50% are like other siblings (dizygotic), but unfortunately I cannot distinguish them.[2] In order to be

[2] Holmlund, Lindahl, and Plug (2008) have also the same limitation in the Swedish data.

included in the sample, both twins must have a partner and have at least one child aged over 22 in 2001.[3]

Education variables are measured in 2001 for children and in 1993 for parents. The levels of education are transformed into years of schooling, according to the maximum level of education obtained, see Table 1.

In Table 2 I present descriptive statistics and in Tables 3 and 4 parameter estimates. Among the covariates, I include the partner's schooling and earnings (measured in 1993) to control for assortative mating, gender and age of the child, and a dummy indicating the first child compared to subsequent ones (Black, Devereux, and Salvanes, 2005b).[4] Including the variable age helps to control for performance of students too young to reach the highest levels of education.[5]

The descriptives in Table 2 show that twin-fathers are, on average, better educated than twin-mothers; twin-mothers' husbands are also better educated and earn more than twin-fathers' wives, as expected. The average age of children is between 30 and 31, while twin-mothers are younger than twin-fathers (55 compared to 58), due to the fact that women are on average younger than men at the birth of their children. We have around 2,000 observations for both samples, which corresponds to 500 pairs of twins. On average, the twins have two children older than 22 years old.

In Table 3 I present results for twin-mothers, while in Table 4 for twin-fathers. I perform cross-section analyses, and then I use the twin-estimator to control for transmitted parental ability. Despite differences in the selection of the samples, results are consistent with the rest of the literature. In the cross section estimations, all the results are in the expected direction: we find a positive effect of parental education, being a woman, and being the first child in the family. In the twin-estimations, mothers' education loses its significance, while the impact of father's schooling is still significantly different from zero. Similar to the research cited above, the effect of mother's education surprisingly does not differ significantly from zero.

How much information can the twin-estimator exploit? In Table 5, I report the total variance, the within variance, and the percentage of twins with the same level of schooling. In the twin-fathers' sample around 50% of twins have the same level of education, while around 60% in the twin-mothers' sample. This implies that the twin-estimator only exploits information from less than

[3] Holmlund, Lindahl, and Plug (2008) selected children aged over 22, Antonovics and Goldberger (2005) from 18 years old, Behrman and Rosenzweig (2002) include also children younger than 18.

[4] Behrman and Rosenzweig (2002) include controls for partner's schooling and earnings. Holmlund, Lindahl and Plug (2008) also include gender of the child.

[5] De Haan and Plug (2006) propose and compare different methodologies to treat these censored observations, and conclude that using parental expectations if they were realizations seems to deal better with censoring problems. Register data do not provide this kind of information.

half of the original samples. To what extent does small sample size affect the estimates? Can the lower level of female variability contribute to the explanation of different results between genders?

4. Small Sample Size and Low Variability in Parental Education

The aim of this section is to understand how small sample sizes and low variability in parental education, typical of twins' samples, affect the power of the test when using the twin-estimator. The idea is to work through simulations. I would like to have a "large perfect population": "perfect" means without problems of ability bias, "large" means large enough to draw samples from it. Then I could estimate the "true" effect of parental education on children's education. Suppose that there were an effect. Then, I could draw small samples, estimate the effect, and count how many times I would reject the null hypothesis. This would measure the power of the test.

We do not know the true effect, and we do not have a population for estimating it. But we can pretend that the true effect is the one obtained using a sample of siblings. The sibling-estimator has often been used for the same purpose as the twin-estimator: to try to separate the effect of interest from the idiosyncratic characteristics of the family (Behrman and Wolfe, 1989; Neumark and Korenman, 1994; Altonji and Dunn, 1996; Ashenfelter and Zimmerman, 1997; Aaronson, 1998; Ermisch and Francesconi, 2000). Using siblings instead of twins has some advantages but also evident shortcomings. On the one hand, they all have the same family background, experience similar lifetime events (but at different times), are more representative of the population, and can provide larger samples and more precise estimations. On the other hand, they do not share the same genes and the results obtained are then potentially biased. Before proceeding with simulations, I show the results using the sibling-estimator, and using siblings whose age difference is, at most, 24/36/48/60 months. I also show results of the sibling-estimator when I include in the sample all siblings (regardless of the difference in age between them). Finally, I consider the results obtained from all children (cross-sectional analysis) that can be suggestive of the importance of the ability bias. The estimated effects are reported in Tables 6, 7 and 8.

In Table 6, we can observe the estimated effects of mother's schooling. In the first row, I report the result obtained with the twin-estimator. When I use the sibling-estimators, the effect is around 0.105, slightly larger when I consider all siblings, regardless of their age. The effect estimated for all children is substantially larger, showing the importance of the "family effect" in this context.

In Table 7, we can observe the effects estimated of father's schooling. Surprisingly, the effect obtained with the twin-estimator is larger in size than the one obtained with the sibling-estimator. The larger estimate in the case of twins may be the consequence of the small sample size.

When I use the sibling-estimators, the effect is around 0.125. The cross-sectional effect is still larger, as it was for mothers.

There are two interesting facts, when comparing results across mothers and fathers. First, the mothers' effect, on average, seems smaller than father's effect. However, the difference is not statically significant. Second, in the mother's case, the bias seems to be more important. We go from 0.103-0.115 (sibling-estimators) to 0.222 (cross-section) for mothers, while from 0.124-0.130 (sibling-estimators) to 0.211 (cross-section) for fathers. The difference between the two cross-section estimates is instead significantly different from zero. The fact that ability is transmitted more between mothers and children than between fathers and children can be explained by the fact that mothers traditionally spend more time with their children.[6]

Finally, I show the variability in the samples (Table 8). All variability measures are increased when considering siblings instead of twins, as expected. Mothers' education has less variability than fathers' education in all samples.

For the simulations, I choose as my "large perfect population" the sample of all siblings. Using this, the "true" effect of mother's schooling is 0.115 and the "true" effect of father's schooling is 0.126. We need to keep in mind that these estimates set upper limits, especially in the mothers' case since the ability bias seems to be more of an issue.

I first concentrate on the size of the twins' samples: I have 500 pairs of female twins, and 503 of male twins. I draw 1,000 samples of these sizes from all female siblings and all male siblings, and I estimate the effect of parental education. Since the passage from 25,000 observations (pairs of siblings) to 500 (pairs of twins) is very big, I also consider an intermediate step with a sample size of around 1,000 pairs of siblings. In Tables 9 and 10, I summarize the results, obtained with samples of around 1,000 and 500 pairs of siblings, for mothers and fathers respectively.

In the first rows I show the mean of the estimated effect and the mean of its standard error. Then the percentage of times the null hypothesis has been rejected[7] when the type I error (or alpha) is set at 10%, 5%, and 1%.[8] In the remaining rows, some characteristics of the drawn samples are shown: number of pairs of siblings, number of siblings, total and within variance of parental education, and percentage of parents with the same years of schooling. If we look at the power of the tests, we can see the effect of small sample size on the probability of finding a significant effect of parental education. With 1,000 pairs of siblings (Table 9), we observe that the power of the test is

---

[6] Björklund, Lindahl, and Plug (2006) also find that mother's pre-birth factors contribute more to the intergenerational transmission of education than father's pre-birth factors.

[7] Counting how many times the null hypothesis is rejected, given that the null hypothesis is false, corresponds to the complement of the type II error.

[8] Alpha or type I error is the probability of rejecting the null hypothesis when the null hypothesis is true.

approximately 1 when estimating the effect of father's education, while somewhat lower when estimating the effect of mother's education. With only 500 pairs of siblings (Table 10), which corresponds to my actual twins' sample size, differences in the power of the test between father and mother's education effect widen: when alpha is set equal to 0.05, we are about 20 percentage points more likely to reject the null hypothesis of zero effect for fathers than for mothers.

However, the samples used for the simulations, which are randomly selected, have more within variability than the actual twins' sample (Table 8). Also the percentage of siblings with the same years of schooling is lower. The percentage of twin-mothers equally educated is 59.8 while the percentage of sibling-mothers equally educated is 48.2 (within variance rising from 0.74 to 1.11). The same for fathers: the percentage of equally educated falls from 49.5 (twins) to 36.4 (siblings) while the within variance rises from 1.18 (twins) to 1.71 (siblings). This is the obvious consequence of working with less similar individuals. However, the within variability can be very important for my measure of the power of the test.

In order to reduce the within variability in my simulations, I first calculate the variance in parental education within each family of twins and siblings. Then, I match every twin family with variance *v* with one sibling family randomly selected among the ones with the same variance *v*. Finally, I drop all information about twins and I end up with a sample of siblings families with the same number of families as in my original twins' sample and with the same within variability. Results for mothers and father are reported in the first two columns of Table 11. The power of the test is generally lower, especially for mothers and for a lower level of alpha. The differences between fathers and mother have also expanded. Given a positive effect of parental education, we are likely to find a result significantly different from zero 8 times out of 10 for fathers, and for less than 6 times out of 10 for mothers (with alpha equal to 0.05).

Finally, in the final group of simulations (4^{th} column of Table 11), I look at the power of the test using the sibling-mothers but with the twin-fathers' variability. I match every twin-father family with variance *v* with one sibling-mother family randomly selected among the ones with same variance *v*. The idea is to measure the power of the test when we estimate the mother's effect, if we had the same "quality" of information we have for fathers. The probability of rejecting the null hypothesis has increased, but not to the same level as fathers: for example, with alpha equal to 0.05, the power would go from 0.571 to 0.674 for mothers, while it is 0.802 for fathers.

For all of these simulations, we need to remember that results are to be considered as upper limits, since ability is not taken into account in the same way as for twins. To what extent can these results be generalized? It is difficult to say. Previous studies have in common with my study a lower variation in maternal than in paternal education but, in general, they seem to have larger within variability than my Norwegian sample.[9] My empirical exercises can provide hints for investigation with other sources of data.

From a statistical point of view, we have observed that small sample size and low within variability have an impact on the probability of rejecting the hypothesis that the parameter of interest is zero. This is particularly true for mothers: within variability is particularly low for women, and seems to be the consequence of schooling decisions of previous generations. Sisters seem more likely to study the same number of years.

From a substantive point of view, however, we have evidence of a smaller effect of mother's education. In the existing literature, a positive and significant effect of mother's schooling has been found when using the reform of compulsory schooling as an instrument, and selecting less educated women (Black, Devereux, and Salvanes, 2005a; Holmlund, Lindahl, and Plug, 2008). To explore this issue in more detail, in the next paragraph, I examine the impact of one additional year of education at the bottom and at the top of the educational distribution.

5. Use of Not Random Samples of Parents

In order to study heterogeneous effects of an increase in schooling along the educational distribution, I divide the samples of twins and siblings into two parts: one part where both twins/siblings have fewer than 11 years of education (bottom part) and another part where they both have 12 years or more (top part). Table 12 summarizes the results. Using the twin-estimator, I find a positive and marginally significant effect of mother's schooling in the bottom part; while a positive and strong effect of father's education only in the top part. When we look at siblings, these results are confirmed: we observe a positive effect of mother's schooling, but relatively small in the top part of the distribution. The effect of father's schooling is always positive but smaller than the effect of mother's in the bottom part of the distribution.

Using the same simulation setting, we measure the power of the test using a small sample and low variability in parental education (500 families, around 2,000 children). The bottom part of the education distribution has the same small amount of variation for mothers and fathers, which may explain the low power of the test in both cases (Table 13). The power of the test is extremely low for the effect of an increase in education among highly educated mothers, assuming that there is any effect (Table 14). On the other hand, we are very likely to get a significant result when looking at highly educated fathers (Table 14).

[9] For example, Holmlund, Lindahl, and Plug (2008) have a sample of twins where 61% of twin-pairs have different level of education.

How can we interpret these results? They seem to confirm those obtained using different strategies than twins. Studies which make use of compulsory schooling reform as an instrument (Black, Devereux, and Salvanes, 2005a; Holmlund, Lindahl, and Plug, 2008) find a positive effect of education only for low educated mothers. In this paper, using the twin-estimator, we observe a positive effect for mothers[10] and not for fathers (Table 12).

Studies which make use of adoption are likely to use samples of parents better educated than the average ones. They find a positive effect of father's schooling, and no effect of mother's (Holmlund, Lindahl, and Plug, 2008; Plug, 2004; Björklund, Lindahl, and Plug, 2006). In my analysis, concentrating on the top part of the educational distribution, I find a small effect of mother's education (and very low likelihood to observe it significant) and a positive effect of father's education (very likely to be observed significant).

The part of the educational distribution used in each analysis seems a very most important factor when explaining different results by gender. This finding helps to reconcile the (conflicting) results in the literature when using different identification strategies.

Many substantial questions need still to be answered: why does increasing education of highly educated women not increase the schooling of the next generation? Are these women working more and spending less time with their children? And why is father's effect negligible in the bottom part of the distribution? Does assortative mating play any role? This calls for further investigation.

6. Redefining the outcome variable "Years of Schooling"

In the data, in addition to the maximum level of education, there are two other pieces of information that could be used to define the years of schooling variables: the age at which education was completed and the level of education currently attended.[11]

I now make use of these variables to examine the robustness of my previous results. After having attributed the years of schooling corresponding to the maximum level obtained, I allow for the possibility that a person stops her/his education between two levels. For example, a person states 4 as maximum level which would imply s/he has studied until the age of 19. The corresponding years of schooling would be 12. If the age of schooling completion in the data is 21, I update her years of schooling to 14. See Table 15 for a comprehensive definition of the variable. This adjustment procedure has been applied to children with completed education before the first available year of the

[10] The effect is significant but only at the 10% level. However, in Section 6, where I exploit more information to define years of schooling, the effect of the mother, in the bottom part of the distribution, becomes larger and more significant (0.243**).

[11] They have not been used so far in order to have a definition of the variables closer to that used by Holmlund, Lindahl and Plug (2008).

panel (1993). For children still enrolled in school during the panel (1993-2001), I can observe whether they attended any further year of schooling compared to the maximum qualification of education stated.[12]

The comparison between the results with the new definition and the results with the old definition can also be suggestive of the downwards bias due to measurement error: not observing the actual length of the schooling but transforming the levels in years is indeed an example of measurement error.

I report results obtained using the twin-estimator, also separately for low and high educated parents; and the respective simulations to measure the power of the test. In parenthesis, I will show the corresponding result with the "old" definition of the variables, in order to make the comparison easier.

When estimating with samples of twins, I obtain 0.085 (0.053) for mothers and 0.209*** (0.192***) for fathers. The probability to reject the null hypothesis, with alpha set to 0.05, is 0.699 (0.571) for mothers and 0.893 (0.802) for fathers. If mothers had the same variability in their education as fathers, the power of the test would increase to 0.813 (0.674).

When analyzing low educated and high educated twin-parents separately, I find that the effect is 0.243** (0.169*) for low educated mothers and 0.207 (0.139) for low educated fathers. The respective powers of the test are 0.643 (0.604) and 0.661 (0.600) with alpha set to 0.05. The effect of high educated mothers' education is 0.145 (0.119) while the effect of high educated fathers' education is 0.258*** (0.333***). The respective powers of the test are 0.540 (0.243) and 0.894 (0.867) with alpha set to 0.05.

Estimates of the intergenerational transmission of schooling and the power of the test are larger than in the previous analyses, as expected, because of the increased variability in both parental and children's education. However, the direction and the statistical significance of the results and differences between mother's and father's effects and power of the tests have not changed.

7. Conclusions

The aim of the paper is to look at the effect of parental education on children's schooling, with particular attention to the difference between the mother's and the father's effects from a statistical point of view. In this paper, using the twin and the sibling-estimator, my results confirm the previous ones which were obtained with different identification strategies (adopted children, schooling reform) and different sources of data: a positive effect of mother's schooling only at the bottom part of the

[12] Suppose the maximum qualification stated in 2001 is the fourth level (secondary school, 12 years of schooling). I can observe whether they attended University any year between 1993 and 2001, even without obtaining a degree.

educational distribution, and a positive strong effect of father's schooling at the top. The part of the educational distribution involved in any identification strategy seems to be a key aspect to take into account to reconcile the (conflicting) results found in the literature.

The positive effect of an increase in education for lower educated mothers is very important from a policy point of view: in contexts where the average level of schooling is low, increasing women's education may still have beneficial effects for children's schooling. What remains unexplained is why the mother's effect is negligible in the top part of the distribution and, on the other hand, why father's effect is weaker in the bottom part of the distribution. Further research is needed. Another contribution of this paper is to assess to what extent small sample sizes and low variation in parental education affects the power of the test when implementing the twin- estimator. Assuming an effect for both mothers and fathers, I demonstrate that we are more likely to observe it for fathers than for mothers. This suggests the importance of considering statistical issues when comparing results for different sub-groups of the population.

Another incidental contribution is to provide the opportunity to compare results obtained employing the twin-estimator and the sibling-estimator. Despite the shortcomings in using siblings instead of twins, estimates show more precision and stability.

Tables

Table 1: Level of Education and Duration in Years in the Norwegian Educational System

Level of education	Duration	Cumulative years of schooling
1	3	3
2	6	9
3	2	11
4[a]	1	12
5	1	13
6	4	16
7	2	18
8	3	21

Notes: [a] After the upper secondary school (level 4), students can decide to go to the University (level 6) or to complete their education with another year of study (level 5) which cannot be classified neither as secondary nor tertiary education. After level 5 they can still have access to level 6 (Norwegian Standard Classification of Education, Revised 2000). Source: Norwegian Standard Classification of Education, Revised (2000).

Table 2: Descriptive Statistics of the Samples of Twin-mothers and Twin-fathers

	Twin-mothers		Twin-fathers	
	Mean	Std. dev.	Mean	Std. dev.
Child's schooling	12.91	2.31	12.97	2.24
Twin's schooling	10.71	1.99	11.36	2.43
Partner's schooling	11.53	2.59	10.91	2.01
Partner's earnings	2.23	1.71	1.12	0.78
Child is a woman	0.48		0.48	
First child	0.50		0.49	
Child's age	30.4	6.5	30.7	6.4
Twin's age	55.1	7.4	58.2	8.2
Observations	1,983 (500 families)		1,996 (503 families)	

Notes: descriptive statistics of the variables related to the twin, his/her partner, and his/her child. Earnings are expressed in NOK and divided by 100,000.

Table 3: Effect of Mother's Years of Schooling on Child's Years of Schooling

	Cross-section		Twin-estimator	
	Beta	St err	Beta	St err
Twins' schooling	0.193***	0.029	0.053	0.051
Partner's schooling	0.192***	0.023	0.072**	0.034
Partner's earnings	0.029	0.032	0.016	0.047
Child is a woman	0.358***	0.097	0.370***	0.103
Child's age	0.003	0.008	-0.014	0.019
First child	0.222**	0.098	0.258**	0.115
Constant	8.201***	0.433	11.594***	0.858
Observations	1,983 (500 families)			

Table 4: Effect of Father's Years of Schooling on Child's Years of Schooling

	Cross-section		Twin-estimator	
	Beta	St err	Beta	St err
Twins' schooling	0.244***	0.022	0.192***	0.039
Partner's schooling	0.110***	0.028	0.044	0.039
Partner's earnings	0.341***	0.066	0.257***	0.098
Child is a woman	0.400***	0.093	0.488***	0.098
Child's age	0.016**	0.008	-0.039**	0.016
First child	0.243***	0.093	0.483***	0.109
Constant	7.801***	0.417	10.741***	0.752
Observations	1,996 (503 families)			

Table 5: Variance of Parental Schooling in the Samples of Twin-mothers and Twin-fathers

	Total variance	Within variance	Same number of years of schooling
Twin-mothers	3.97	0.74	59.8 %
Twin-fathers	5.92	1.18	49.5 %

Table 6: Effect of Mother's Years of Schooling on Child's Years of Schooling, using the Sibling-estimator and Different Samples of Siblings

Relationship	Beta	Std. err.	Families	Observations
Twins	0.053	0.051	500	1,983
Siblings (24 m)	0.103***	0.014	5,294	21,749
Siblings (36 m)	0.110***	0.010	9,577	40,138
Siblings (48 m)	0.105***	0.008	13,059	55,159
Siblings (60 m)	0.103***	0.008	15,723	66,774
All Siblings	0.115***	0.006	23,127	100,621
All Children	0.222***	0.002	-	431,848

Notes: The numbers of months, in brackets, indicate the maximum distance between the births of the pair of siblings. In the regressions, I also control for partner's schooling and earnings; child's age, gender and first-born child.

Table 7: Effect of Father's Years of Schooling on Child's Years of Schooling, using the Sibling-estimator and Different Samples of Siblings

Relationship	Beta	Std. err.	Families	Observations
Twins	0.192***	0.039	503	1,996
Siblings (24 m)	0.130***	0.010	5,298	25,576
Siblings (36 m)	0.124***	0.008	10,729	46,819
Siblings (48 m)	0.125***	0.006	14,711	64,659
Siblings (60 m)	0.126***	0.006	17,785	78,517
All Siblings	0.126***	0.005	26,910	121,512
All Children	0.211***	0.001	-	476,463

Notes: The numbers of months, in brackets, indicate the maximum distance between the births of the pair of siblings. In the regressions, I also control for partner's schooling and earnings; child's age, gender and first-born child.

Table 8: Variance of Parental Schooling in the Samples of Siblings

Relationship	Mothers			Fathers		
	Total variance	Within variance	Same number of years of schooling	Total variance	Within variance	Same number of years of schooling
Twins	3.97	0.74	59.8	5.95	1.18	49.5
Siblings (24 m)	4.15	1.04	50.9	6.11	1.55	39.4
Siblings (36 m)	4.26	1.09	49.5	6.20	1.58	38.1
Siblings (48 m)	4.24	1.11	48.9	6.29	1.65	37.2
Siblings (60 m)	4.19	1.11	48.5	6.26	1.66	37.1
All Siblings	3.99	1.10	48.2	6.05	1.71	36.4
All Children	4.34	-	-	6.39	-	-

Notes: The numbers of months, in brackets, indicate the maximum distance between the births of the pair of siblings.

Table 9: Power of the Test, Drawing 1,000 Families

	Mothers		Fathers	
	Mean	Std. dev.	Mean	Std. dev.
Beta	0.111	0.033	0.126	0.027
Std error	0.030	0.001	0.024	0.001
Power of the test (alpha = 0.10)	0.966		0.999	
Power of the test (alpha = 0.05)	0.937		0.999	
Power of the test (alpha = 0.01)	0.851		0.993	
Total variance	3.99	0.20	6.07	0.27
Within variance	1.10	0.07	1.71	0.10
Same schooling	48.3		36.3	
Observations	4,351	62	4,542	62

Notes: Mean of the effect of parental schooling on child’s years of schooling (controlling for partner’s education and earnings; child’s age, gender and order of birth), power of test (for different levels of alpha - type I error) and variance of parental education, drawing 1,000 families (1,006 for fathers) from the sample of all siblings (1,000 simulations).

Table 10: Power of the Test, Drawing 500 Families

	Mothers		Fathers	
	Mean	Std. dev.	Mean	Std. dev.
Beta	0.115	0.05	0.124	0.039
Std error	0.043	0.002	0.034	0.001
Power of the test (alpha = 0.10)	0.806		0.962	
Power of the test (alpha = 0.05)	0.740		0.935	
Power of the test (alpha = 0.01)	0.545		0.832	
Total variance	3.97	0.28	6.06	0.38
Within variance	1.10	0.10	1.71	0.14
Same schooling	48.3		36.2	
Observations	2,178	43	2,273	44

Notes: Mean of the effect of parental schooling on child's years of schooling (controlling for partner's education and earnings; child's age, gender and order of birth), power of test (for different levels of alpha - type I error) and variance of parental education, drawing 500 families (503 for fathers) from the sample of all siblings (1,000 simulations).

Table 11: Power of the Test, Drawing 500 Families with Low Variability in Parental Schooling

	Mothers		Fathers		Mothers with fathers' variability	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Beta	0.113	0.062	0.120	0.048	0.109	0.052
Std error	0.052	0.001	0.041	0.001	0.044	0.001
Power of the test (alpha = 0.10)	0.672		0.870		0.773	
Power of the test (alpha = 0.05)	0.571		0.802		0.674	
Power of the test (alpha = 0.01)	0.367		0.596		0.481	
Total variance	3.65	0.25	5.51	0.37	4.17	0.23
Within variance	0.73	0.02	1.12	0.03	1.09	0.03
Same schooling	59.8		49.5		49.6	
Observations	2,125	42	2,203	40	2,115	40

Notes: Mean of the effect of parental schooling on child's years of schooling (controlling for partner's education and earnings; child's age, gender and order of birth), power of test (for different levels of alpha - type I error) and variance of parental education, drawing not randomly 500 families (503 for fathers) from the sample of all siblings (1,000 simulations).

Table 12: Effect of Parental Schooling on Child's Years of Schooling at the Top and at the Bottom of the Educational Distribution

	Both twins/siblings <= 11 years of schooling		Both twins/siblings >= 12 years of schooling	
	Mothers	Fathers	Mothers	Fathers
Twins				
Beta	0.169*	0.139	0.119	0.333***
Std error	0.097	0.128	0.132	0.093
Families	371	242	52	111
Observations	1,520	1,005	184	388
Siblings				
Beta	0.163***	0.153***	0.053***	0.114***
Std error	0.012	0.013	0.019	0.011
Families	16,608	13,609	2,364	5,465
Observations	75,424	65,636	8,490	20,481

Notes: In the regressions, I also control for partner's schooling and earnings; child's age, gender and first-born child.

Table 13: Power of the Test, Drawing 500 Families with Low Variability in Parental Schooling (Bottom Part of the Educational Distribution)

	Both twins/siblings <= 11 years of schooling			
	Mothers		Fathers	
	Mean	Std. dev.	Mean	Std. dev.
Beta	0.164	0.083	0.151	0.079
Std error	0.072	0.003	0.067	0.003
Power of the test (alpha = 0.10)	0.702		0.701	
Power of the test (alpha = 0.05)	0.604		0.600	
Power of the test (alpha = 0.01)	0.404		0.384	
Total variance	1.00	0.01	0.99	0.01
Within variance	0.34	0.02	0.35	0.02
Same schooling	63.5		62.3	
Observations	2,269	43	2,426	48

Notes: Mean of the effect of parental schooling on child's years of schooling (controlling for partner's education and earnings; child's age, gender and order of birth), power of test (for different levels of alpha - type I error) and variance of parental education, drawing not randomly 500 families (503 for fathers) from the sample of all siblings (1,000 simulations).

Table 14: Power of the Test, Drawing 500 Families with Low Variability in Parental Schooling (Top Part of the Educational Distribution)

	Both twins/siblings >= 12 years of schooling			
	Mothers		Fathers	
	Mean	Std. dev.	Mean	Std. dev.
Beta	0.054	0.041	0.115	0.039
Std error	0.042	0.001	0.037	0.001
Power of the test (alpha = 0.10)	0.340		0.918	
Power of the test (alpha = 0.05)	0.243		0.867	
Power of the test (alpha = 0.01)	0.091		0.695	
Total variance	3.79	0.09	6.14	0.19
Within variance	1.54	0.08	1.85	0.12
Same schooling	45.9		42.9	
Observations	1,796	26	1,918	31

Notes: Mean of the effect of parental schooling on child's years of schooling (controlling for partner's education and earnings; child's age, gender and order of birth), power of test (for different levels of alpha - type I error) and variance of parental education, drawing not randomly 500 families (503 for fathers) from the sample of all siblings (1,000 simulations).

Table 15: New Definition of Years of Schooling

Level of Education	Correspondent Years of Schooling	Theoretical Age -end	Actual Ending Age between	Imputed Years of Schooling
1	3	10	11-15	4-9
2	9	16	17	10
3	11	18	-	-
4	12	19	20-22	13-15
5	13	20	21-23	14-16
6	16	23	24	17
7	18	25	26-27	19-20
8	21	28	-	-

Example: a person states 4 as maximum level, which would imply s/he has studied until the age of 19. The correspondent years of schooling would be 12. If the age of schooling completion in the data is 21, 14 years of schooling are imputed.

References

Aaronson, D. (1998): Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes. *The Journal of Human Resources*, **33**, 915-946.

Altonji, J.G. and T.A. Dunn (1996): Using Siblings to Estimate the Effect of School Quality on Wages. *The Review of Economics and Statistics*, **78**, 665-671.

Antonovics, K.L. and A.S. Goldberger (2005): Does Increasing Women's Schooling Raise the Schooling of the Next Generation? Comment. *The American Economic Review*, **95**, 1738-1744.

Ashenfelter, O. and A. Krueger (1994): Estimates of the Economic Return to Schooling from a New Sample of Twins. *The American Economic Review*, **84**, 1157-1173.

Ashenfelter, O. and D.J. Zimmerman (1997): Estimates of the Returns to Schooling from Sibling Data: Fathers, Sons, and Brothers. *The Review of Economics and Statistics*, **79**, 1-9.

Behrman, J.R. and M.R. Rosenzweig (2002): Does Increasing Women's Schooling Raise the Schooling of the Next Generation? *The American Economic Review*, **92**, 323-334.

Behrman, J.R. and B.L. Wolfe (1989): Does More Schooling Make Women Better Nourished and Healthier? Adult Sibling Random and Fixed Effects for Nicaragua. *The Journal of Human Resources*, **24**, 644-663.

Björklund, A., M. Lindahl and E. Plug (2006): The Origins of Intergenerational Associations: Lessons from Swedish Data. *The Quarterly Journal of Economics,* **121**, 999-1028.

Black, S.E., P.J. Devereux and K.G. Salvanes (2005a): Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital. *The American Economic Review*, **95**, 437-449.

Black, S.E., P.J. Devereux and K.G. Salvanes (2005b): The More The Merrier? The Effect of Family Size and Birth Order on Children's Education. *The Quarterly Journal of Economics,* **120**, 669-700.

Bound, J. and G. Solon (1999): Double Trouble: On the Value of Twin-Based Estimation of the Return to Schooling. *Economics of Education Review*, **18**, 169-182.

De Haan, M. and E. Plug (2006): Estimates of the Effect of Parent's Schooling on Children's Schooling Using Censored and Uncensored Samples. IZA Discussion Paper series, no 2416.

Ermisch, J. and M. Francesconi (2000): Educational Choice, Families, and Young People's Earnings. *The Journal of Human Resources*, **35**, 143-176.

Griliches, Z. (1979): Siblings Models and Data in Economics: Beginning of a Survey. *The Journal of Political Economy*, **87**, S37-S64.

Haveman, R. and B. Wolfe (1995): The Determinants of Children's Attainments. *Journal of Economic Literature*, **33**, 1829-1878.

Holmlund, H., M. Lindahl and E. Plug (2008): Estimating Intergenerational Schooling Effect: A Comparison of Methods. IZA Discussion Paper, no 3630.

Light, A. and A. Flores-Lagunes (2006): Measurement Error in Schooling: Evidence from Samples of Siblings and Identical Twins. *Contributions to Economic Analysis and Policy*, Volume 5, Issue 1, Article 14.

Mittler, P. (1971): *The Study of Twins.* Penguin Books, Baltimore.

Mowrer, E.R. (1954): Some Factors in the Affectional Adjustment of Twins. *American Sociological Review*, **19**, 468-471.

Neumark, D. (1994): Biases in Twin Estimates of the Return to Schooling: A Note On Recent Research. National Bureau of Economic Research, Technical Working Paper No. 158.

Neumark, D. and S. Korenman (1994): Sources of Bias in Women's Wage Equations: Results Using Sibling Data. *The Journal of Human Resources*, **29**, 379-405.

Norwegian Standard Classification of Education, Revised (2000), Official Statistics of Norway, C 751, Statistics Norway.

Plug, E. (2004): Estimating the Effect of Mother's Schooling on Children's Education Using a Sample of Adoptees. *The American Economic Review*, **94**, 358-368.

Schieve, L.A. et al. (2002): Low and Very Low Birth Weight in Infants Conceived with Use of Assisted Reproductive Technology. *New England Journal of Medicine*, **346**, 731-737.

Stewart, E.A. (2000): *Exploring Twins: Towards a Social Analysis of Twinship.* McMillan Press, London.